

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

1c970 U.S. PTO  
09/942905  
08/31/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日  
Date of Application:

2001年 5月31日

出 願 番 号  
Application Number:

特願2001-165075

出 願 人  
Applicant(s):

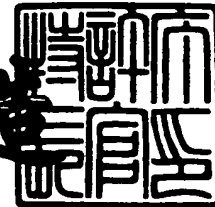
株式会社日立製作所

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 8月10日

特 許 庁 長 官  
Commissioner,  
Japan Patent Office

及 川 耕 造



【書類名】 特許願

【整理番号】 H100677

【提出日】 平成13年 5月31日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 17/30

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社  
日立製作所 中央研究所内

【氏名】 今一 修

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社  
日立製作所 中央研究所内

【氏名】 久光 徹

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社  
日立製作所 中央研究所内

【氏名】 岩山 真

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社  
日立製作所 中央研究所内

【氏名】 西岡 真吾

【発明者】

【住所又は居所】 東京都国分寺市東恋ヶ窪一丁目 2 8 0 番地 株式会社  
日立製作所 中央研究所内

【氏名】 丹羽 芳樹

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社 日立製作所

【代理人】

【識別番号】 100091096

【弁理士】

【氏名又は名称】 平木 祐輔

【手数料の表示】

【予納台帳番号】 015244

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 文書検索システムおよびサーバ

【特許請求の範囲】

【請求項 1】 文書データベースの検索結果をもとに、複数の文書データベースの中から次に検索する文書データベースを指定して文書検索を指示できる文書検索システムにおいて、

文書データベース  $i$  の検索結果をもとに文書データベース  $j$  の検索が行われた回数  $x_{ij}$  を記録した連想検索記録テーブルを保持することを特徴とする文書検索システム。

【請求項 2】 請求項 1 記載の文書検索システムにおいて、前記連想検索記録テーブルを用いて、検索対象となる文書データベースの表示順序を変更することを特徴とする文書検索システム。

【請求項 3】 請求項 1 記載の文書検索システムにおいて、前記連想検索記録テーブルをユーザ別に保持し、前記ユーザ別の連想検索記録テーブルを用い、検索対象となる文書データベースの表示順序をユーザに応じて変更することを特徴とする文書検索システム。

【請求項 4】 請求項 1 記載の文書検索システムにおいて、前記連想検索記録テーブルを用いて、各文書データベースの登録料を計算することを特徴とする文書検索システム。

【請求項 5】 請求項 4 記載の文書検索システムにおいて、文書検索の検索元となった回数と検索先となった回数との和に応じて登録料を計算することを特徴とする文書検索システム。

【請求項 6】 文書データベースの検索結果をもとに複数の文書データベースの中から次に検索する文書データベースを指定して文書検索を指示できる検索クライアントと複数の文書データベースとの間を仲介するサーバにおいて、

前記検索クライアントからの検索要求を解析する検索要求解析手段と、

前記検索クライアントによって指定された文書データベースに前記検索要求解析手段で解析した検索要求を送る検索要求発行手段と、

前記指定された文書データベースによる検索結果を前記検索クライアントに送

信する手段と、

文書データベース  $i$  の検索結果をもとに文書データベース  $j$  の検索が行われた回数  $x_{ij}$  を記録した連想検索記録テーブルを保持する連想検索記録テーブル保持手段とを備えることを特徴とするサーバ。

【請求項 7】 請求項 6 記載のサーバにおいて、前記連想検索記録テーブルを用いて前記検索クライアントに表示する検索対象となる文書データベースの表示順序を変更する表示順序変更手段を備えることを特徴とするサーバ。

【請求項 8】 請求項 6 記載のサーバにおいて、前記連想検索記録テーブル保持手段はユーザ別の連想検索記録テーブルを保持し、前記ユーザ別の連想検索記録テーブルを用い、ユーザに応じて前記検索クライアントに表示する検索対象となる文書データベースの表示順序を変更する表示順序変更手段を備えることを特徴とするサーバ。

【請求項 9】 請求項 6 記載のサーバにおいて、前記連想検索記録テーブル保持手段に保持された前記連想検索記録テーブルを用いて各文書データベースの登録料を計算することを特徴とするサーバ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数の文書データベースを関連付けて連想検索できる文書検索システムおよびサーバに関する。

【0002】

【従来の技術】

文書情報の電子化が進むにつれ、複数の異なる種類の文書データベースを対象とする高精度な文書検索技術が求められている。また、複数の文書データベース間での文書群の関連性を調べたいという要求も高まっている。例えば、興味のある新聞記事に対して、それに関連する百科事典の項目を検索したいという要求は多い。

【0003】

現在実用となっているキーワード検索では、複数の文書データベースを切り換

えて検索することは可能であるが、ある文書データベースの文書群に対して、それに関連する文書群を、同一文書データベース、あるいは、別の文書データベースから検索すること（文書連想検索と呼ばれる検索方式）は不可能である。

【0004】

同一文書データベース内に限れば、文書間の関連度をあらかじめ計算しておくことで、文書群を検索入力とした文書連想検索を実現することはできる。しかし、複数の文書データベースに対しては、あらかじめ計算すべき文書間の関連度が文書データベース数の増加に伴って組み合わせ数が爆発的に増加するため、現実的には不可能となっている。

【0005】

これに対して、特開2000-155758号公報「複数文書データベースを対象とした文書検索方法及び文書検索サービス」には、利用者が指定した文書データベース中の任意の文書群に対して、その文書群に関連する文書群を任意の文書データベースから効率良く検索する方法が開示されている。この方法では、文書群として入力された検索入力内の特徴的な単語のみを概要として検索に使うことで、高速な文書連想検索を実現している。この方法により、利用者は複数の異なる種類の文書データベースを切り換えながら、文書群の関連性を調べることで、高精度かつ効率の良い文書検索を行なうことが可能となる。

【0006】

【発明が解決しようとする課題】

複数の文書データベースを関連付けて連想検索できる文書検索サーバの検索クライアントは、検索対象となる文書データベースを選択するための機構を有している。利用者は検索したい文書データベースを文書データベースリストから選択することで検索を行なうわけであるが、一般に、選択可能な文書データベースリストの表示順序は固定されている。利用者が検索しようとするデータベースがリストの下の方に表示されていた場合は、選択の手間がかかるため、利用者にとっての利便性が低いものとなっていた。利用者の利便性を高めるには、利用者の利用状況に応じて文書データベースの順序付けを変える必要がある。

【0007】

文書データベースが登録されている検索サーバは、一般に、文書データベースの所有者から登録料と手数料を徴収する。登録料は定額料金あるいは文書データベースのアクセス数に応じた料金であることが多い。手数料は、文書データベースの所有者が得た利益の一部を徴収するもので、一般に、利用者が文書データベースの所有者に支払う購読料の一部が手数料として支払われる。

登録料や手数料を計算するためには、文書データベースの評価を行なう必要がある。従来のキーワード検索サーバでは、文書データベースのアクセス数を使うのが一般的であったが、連想検索サーバでは、アクセス数だけでは評価できない要因がある。例えば、購読されることの多い文書データベースの登録料は下げ、少ない文書データベースの登録料は上げるとすると、購読されることは少ないが、連想検索の元記事としては頻繁に使われる文書データベースがあった場合、この文書データベースに対する評価は低いものとなり、登録料が上げられることになる。しかし、この文書データベースは連想検索の元記事として使われることで、連想検索サーバ全体の活性度を高めるのに貢献しており、そのことは高く評価されるべきである。つまり、アクセス数だけでは文書データベースの評価には不十分であり、連想検索サーバ全体の活性度にいかに貢献したかを適切な尺度で評価する必要がある。

本発明では、利用者が検索対象の文書データベースを選択する際の利便性を向上させること、および、文書データベースの適切な評価を行なうこと、を目的とする。

#### 【 0 0 0 8 】

##### 【課題を解決するための手段】

本発明においては、複数の文書データベースを関連付けて連想検索できる文書検索サーバにおいて、連想検索の履歴を連想検索記録テーブルとして記録しておき、それを用いて文書データベースの順序付けを行なう。連想検索記録テーブルは、どの文書データベースからどの文書データベースへ、連想検索が何回行なわれたかを記録するものである。

#### 【 0 0 0 9 】

連想検索記録テーブルを利用することで、検索対象となる文書データベースの

表示順序を検索結果に応じて並べることができる。例えば、ある新聞を検索した結果が表示されている場合には、その新聞からよく連想検索される文書データベースをリストの上位に表示しておく。このように順序付けられた文書データベースリストを用いることで、利用者はより利便性の高いかたちで次に検索しようとする文書データベースを選択することができる。

## 【 0 0 1 0 】

また、文書データベースの登録料や手数料を計算する際にも、連想検索記録テーブルに記録されている、連想元として使われた回数や連想先として使われた回数を利用することで、従来手法よりも適切かつ詳細に登録料や手数料の計算を行なうことが可能となる。

## 【 0 0 1 1 】

すなわち、本発明による文書検索システムは、文書データベースの検索結果をもとに、複数の文書データベースの中から次に検索する文書データベースを指定して文書検索を指示できる文書検索システムにおいて、文書データベース  $i$  の検索結果をもとに文書データベース  $j$  の検索が行われた回数  $x_{ij}$  を記録した連想検索記録テーブルを保持することを特徴とする。

## 【 0 0 1 2 】

連想検索記録テーブルは、検索対象となる文書データベースの表示順序を変更するために利用することができる。連想検索記録テーブルをユーザ別に保持し、ユーザ別の連想検索記録テーブルを用いて、検索対象となる文書データベースの表示順序をユーザに応じて変更するようにしてもよい。

連想検索記録テーブルは、各文書データベースの登録料を計算するために利用することもできる。その際、文書検索の検索元となった回数と検索先となった回数との和に応じて登録料を計算するようにしてもよい。

## 【 0 0 1 3 】

本発明によるサーバは、文書データベースの検索結果をもとに複数の文書データベースの中から次に検索する文書データベースを指定して文書検索を指示できる検索クライアントと複数の文書データベースとの間を仲介するサーバにおいて、検索クライアントからの検索要求を解析する検索要求解析手段と、検索クライ



アントによって指定された文書データベースに検索要求解析手段で解析した検索要求を送る検索要求発行手段と、指定された文書データベースによる検索結果を検索クライアントに送信する手段と、文書データベース  $i$  の検索結果をもとに文書データベース  $j$  の検索が行われた回数  $x_{ij}$  を記録した連想検索記録テーブルを保持する連想検索記録テーブル保持手段とを備えることを特徴とする。

## 【 0 0 1 4 】

## 【発明の実施の形態】

以下、図面を参照して本発明の実施の形態を説明する。以下の図において、同じ機能部分には同一の符号を付し、重複する説明を省略する。

図1は、本発明を実現するためのシステム構成の例を示す概略図である。このシステムは、ユーザが検索要求を入力したり、検索対象となる文書データベースを選択したり、検索結果を表示したりするための検索インタフェースを実現するための検索クライアント20、文書データベースを検索するための検索サーバ40, 50, 60、検索クライアント20と検索サーバ40, 50, 60の仲介をする連想検索サーバ30を通信ネットワーク10により接続して構成される。図に示した例では、文書データベースを検索するための検索サーバとして3つの検索サーバが通信ネットワーク10に接続されているが、通信ネットワークに接続される検索サーバの数は任意である。

## 【 0 0 1 5 】

検索サーバ(40, 50, 60)の備える概略作成手段(401, 501, 601)は、文書データベース(403, 503, 603)から検索された文書群の概略を作成する。ここでの概略とは、文書群の内容をよく表わす単語集合のことである。概略作成手段としては、特開平9-62693号公報「確率モデルによる文書分類方法」などの既存手法が利用できる。

## 【 0 0 1 6 】

一例を示すと、まず、概略を作成しようとする文書群中の全文書を単語に分割して頻度集計する。一般に、ある文書群で良く現れる単語ほどその文書群を代表する度合も高いため、文書群中での出現頻度が高い単語ほど概略に含まれやすいことになる。ただし、「する」などのように、どの文書にも良く現れるような一

般的な単語は概略単語として適当ではない。よって、通常は、文書群が属する文書データベース中での出現頻度も考慮して概略単語を選択する。つまり、指定された文書群での出現頻度が高く、かつ、文書データベース全体での総出現頻度が低い単語ほど、その文書群中でしか現れないという意味で特徴的な単語であり、その文書群を特徴付ける概略として適切である。具体的には、文書群中のそれぞれの単語について、文書群中での出現頻度、文書データベース中での出現頻度を入力とする適当な関数により単語の重みを計算し、ある閾値以上の重みを持つ単語を概略として採用する。

#### 【0017】

検索サーバ(40, 50, 60)の備える検索手段(402, 502, 602)は、連想検索サーバ30から送られてきた検索要求に対して、関連度の高い文書群を文書データベース(403, 503, 603)から検索し、その検索結果を関連度の重み付きで連想検索サーバ30に返す。ここでの検索手段は、例えば、公知のキーワード検索手法により実現できる。つまり、入力である検索要求は重み付きの単語集合であるため、各単語を重み付きの入力キーワードとみなしてOR検索すればよい。その際、検索結果の文書の重み(関連度)は次のように計算できる。検索要求および検索先の文書両方に含まれる各単語について、検索要求中での重みと、検索先の文書における重み(例えば頻度)から総合的な重みを計算し(例えば両重みの積)、さらにそのような単語すべてに関する重みを集計(例えば総和)することで関連度を得る。

検索クライアント20は、検索要求入力手段201、文書データベース選択手段202、検索結果表示手段(文書群指定手段)203、概要表示手段(単語群指定手段)204を備える。

#### 【0018】

図2は検索クライアントの例である。利用者は文書データベース選択ボタン2021により検索対象とする文書データベースを選択し、キーワード入力ボックス2011に検索要求を入力し、キーワード検索指示ボタン2012をクリックすることで検索を行なう。検索結果は検索結果表示手段203によって表示され、検索結果の概要が概要表示手段204によって表示される。検索結果の文書は

閲覧ボタン2031によりその内容を閲覧することができる。この例では、検索対象となる文書データベースとしてDB<sub>1</sub>を選択し、検索要求として「キーワード1」を入力した検索結果を示している。

## 【0019】

検索結果表示手段203は文書群指定手段も兼ねており、文書選択ボタン2032により、任意個の記事を選択し、検索対象とする文書データベースを文書データベース選択手段202により選択し、「文書から検索」ボタン2033をクリックすることで、選択した記事と関連する文書を検索することができる。

## 【0020】

概要表示手段204は単語群指定手段も兼ねており、単語選択ボタン2041により、任意個の単語を選択し、検索対象とする文書データベースを文書データベース選択手段202により選択し、「概略単語から検索」ボタン2042をクリックすることで、概略単語からの検索を行なうことができる。

## 【0021】

連想検索サーバ30は、検索クライアント20から送られてくる検索要求を解析する検索要求解析手段301、検索クライアント20から送られてくる検索対象の文書データベースに応じて、検索サーバ(40, 50, 60)に検索要求を振り分ける機能をもつ検索要求発行手段302、文書群に対する概略単語を検索サーバ(40, 50, 60)に要求する概略単語要求手段303、連想検索の履歴を記録するための連想検索記録テーブルを記録・保持する連想検索記録テーブル保持手段304、文書データベースの表示順序を変更する表示順序変更手段305、文書データベースの登録料を計算する登録料計算手段306を備える。

## 【0022】

検索要求解析手段301は、検索クライアント20から送られてくる検索要求を解析してその中に含まれる単語を同定し検索キーを作成する。検索要求解析手段301は任意のものであってよいが、日本語文に関しては文を単語に分割する形態素解析、英文に関しては単語の原形還元と品詞付けを行なうstemmingを最低含んで構成される。

## 【0023】

検索要求発行手段 3 0 2 は、検索対象となる文書データベースに応じて、検索クライアント 2 0 から送られてくる検索キーを検索要求として検索サーバ (4 0, 5 0, 6 0) に振り分ける。検索キーは、(1) 検索要求解析手段 3 0 1 によって作成された単語集合、(2) 検索クライアント 2 0 が備える検索結果表示手段 (文書群指定手段) 2 0 3 から送られてくる文書 ID、あるいは (3) 検索クライアント 2 0 が備える概要表示手段 (単語群指定手段) 2 0 4 から送られてくる単語集合、である。検索キーが (1) と (3) の場合、その単語集合を検索要求として検索サーバに送る。検索キーが (2) の場合、概略単語要求手段 3 0 3 が検索対象となる文書データベースを備える検索サーバに対して、その文書 ID に対応する文書の概略を要求し、送られてきた概略単語集合を検索要求として検索サーバに送る。

#### 【 0 0 2 4 】

連想検索記録テーブル保持手段 3 0 4 は、連想検索の履歴を連想検索記録テーブルに記録・保持する。図 3 は連想検索記録テーブルの例である。以下、図 3 を参照しながら、連想検索記録テーブルの作成方法を説明する。

例えば、検索対象として利用可能な文書データベースとして、 $DB_1$ ,  $DB_2$ ,  $DB_3$  があるとする。利用者は検索クライアントを利用するにあたって、まずキーワード検索を用いて  $DB_1$  を検索したとする。この場合、連想検索記録テーブルの「keyword」行の「 $DB_1$ 」列である  $x_{41}$  に 1 を加える。次に  $DB_1$  中のある文書を選択し、その文書に関連する文書を  $DB_2$  から検索したとする。この場合、連想検索記録テーブルの「 $DB_1$ 」行の「 $DB_2$ 」列である  $x_{12}$  に 1 を加える。以下、利用者が文書データベースを切り換えながら連想検索を行なった履歴を記録していくことで連想検索記録テーブルの各要素の値を更新していく。

#### 【 0 0 2 5 】

表示順序変更手段 3 0 5 は連想検索記録テーブル保持手段 3 0 4 によって保持されている連想検索記録テーブルをもとに、検索クライアント 2 0 が備える文書データベース選択手段 2 0 2 における文書データベースの表示順序を変更する。順序付けを行なう基準としては、(1) 連想検索元としてよく使われる文書データベースを上位にする、(2) 連想検索先としてよく使われる文書データベース

を上位にする、(3) 連想検索先にも連想検索元にもよく使われる文書データベースを上位にする、(4) 連想検索元の文書データベースを決めた場合の連想検索先としてよく使われる文書データベースを上位にする、が考えられる。

## 【0026】

基準(1)で文書データベースを順序付けるには、連想検索記録テーブル上の各文書データベースに関して行方向に総和をとった値( $X_i: i=1,2,3$ )を比較すればよい。

基準(2)で文書データベースを順序付けるには、連想検索記録テーブル上の各文書データベースに関して列方向に総和をとった値( $Y_j: j=1,2,3$ )を比較すれば良い。

## 【0027】

基準(3)で文書データベースを順序付けるには、連想検索記録テーブル上の各文書データベースに関して行方向の総和と列方向の総和を足し合わせたもの、つまり、基準(1)で得られた値( $X_i: i=1,2,3$ )と基準(2)で得られた値( $Y_j: j=1,2,3$ )を足し合わせたもの( $X_i+Y_j: i=1,2,3$ )を比較すれば良い。

基準(4)で文書データベースを順序付けるには、連想検索記録テーブル上で連想検索元の文書データベースを $DB_i$ に固定した場合に、文書データベース $DB_j$ が連想検索先として使われた回数 $x_{ij}$ を比較すれば良い。

## 【0028】

一般に、利用者がある検索結果を得ている状況で、次に検索しようとする文書データベースの選択は、その時点で得られている検索結果が属する文書データベースに応じて変化するのが自然である。例えば、新聞を検索した後は百科事典がよく検索され、遺伝子データベースを検索した後は生物学辞典がよく検索される、という状況は多く存在する。このような状況では、基準(4)を用いて文書データベースの表示順序を並べかえることで、検索結果の属する文書データベースに応じて、次によく検索される文書データベースが上位に表示されるため、利用者にとってより利便性の高いものとなる。

## 【0029】

表示順序変更手段305によって表示順序が変更される例を図4、図5、図6

を用いて説明する。

図4は検索開始時点の検索クライアント20の例である。検索開始時点では、検索結果表示手段203、概要表示手段204には何も表示されない。検索を始めるにあたり、利用者はキーワード入力ボックス2011に「キーワード1」、検索対象としてDB<sub>1</sub>を選択したする。図4はこの状態を示しており、キーワード検索指示ボタン2012をクリックすることで検索が実行される。

#### 【0030】

今、連想検索記録テーブルとして図5が得られており、表示順序の変更基準として上記基準(4)を用いるとする。この連想検索記録テーブルでは、キーワード検索される文書データベースはDB<sub>1</sub>、DB<sub>2</sub>、DB<sub>3</sub>の順に頻度が高い。よって、表示順序変更手段305は、検索開始時点の文書データベースの表示順序をDB<sub>1</sub>、DB<sub>2</sub>、DB<sub>3</sub>の順にする。一方、DB<sub>1</sub>を検索元とした場合の検索先の文書データベースはDB<sub>3</sub>、DB<sub>2</sub>、DB<sub>1</sub>の順になっている。よって、表示手段変更手段305は、DB<sub>1</sub>を検索した結果における文書データベースの表示順序をDB<sub>3</sub>、DB<sub>2</sub>、DB<sub>1</sub>の順にする。この様子を示したものが図6である。

登録料計算手段306では、連想検索記録テーブルを用いて、検索サーバに文書データベースを登録する際の登録料を計算する。登録料は期間を決めて(例えば1年ごと)徴収するのが一般的である。

#### 【0031】

連想検索記録テーブルを利用することで、文書データベースの評価を行なうことができる。さきに述べたように、文書データベースの評価を行なう際に、連想検索サーバでは、単なるアクセス数だけでは評価できない要因があるが、上記基準(3)を利用することで適切な評価を行なうことができる。アクセス数に相当する上記基準(2)だけでなく、連想検索の元記事として利用される回数として上記基準(1)を加味して評価する上記基準(3)を用いることで、購読されることは少ないが、連想検索の元記事としては頻繁に使われる文書データベースの正当な評価が可能となる。評価の高い文書データベースの登録料を下げ、評価の低い文書データベースの登録料を上げることで、良質な文書データベースが連想検索サーバに集中することになり、利用者にとっても利用価値が高まる。それ

に伴い、利用者の増加も見込まれ、連想検索サーバ自体の活性度も高められるため、良質な文書データベースの所有者が連想検索サーバに文書データベースを登録する価値が高められる。

#### 【0032】

一例として、図3の連想検索記録テーブルを用いたDB<sub>i</sub>に対する登録料の計算式を $n(E-(Z_i-E))$  ( $i=1,2,3$ ) とする。ここで、 $Z_i=X_i+Y_i-x_{ii}$  ( $i=1,2,3$ )、 $E=(Z_1+Z_2+Z_3)/3$  ( $Z_i$ の平均値)、 $n$ は1アクセスあたりの単価である。この計算式を用いると、連想検索サーバが受け取る文書データベースの登録料の総和は、 $n(E-(Z_1-E))+n(E-(Z_2-E))+n(E-(Z_3-E))=3nE=n(Z_1+Z_2+Z_3)$ となる。よって、総登録料を文書データベース全体のアクセス数に応じた料金としながら、評価の高い文書データベースの登録料を下げ、評価の低い文書データベースの登録料を上げることができる。

#### 【0033】

以上の手段を用いることにより、連想検索の履歴を連想検索記録テーブルに保持しながら、検索クライアント20の文書データベース選択手段で表示する文書データベースの表示順序を適切に変更したり、文書データベース毎の登録料を適切に計算したりすることができる。

#### 【0034】

次に図7、図8、図9を用いて、(1)キーワード検索、(2)関連文書検索、(3)概略単語からの検索、における実行の流れを順に説明する。

図7は、検索サーバ40に対してキーワード検索するときの実行の流れを示す図である。まず、利用者は検索クライアント20を用いて、連想検索サーバ30に接続する(T1)。表示順序変更手段305は連想検索記録テーブルを参照して文書データベースの表示順序を変更し、その順序を検索クライアント20に送る(T2)。検索クライアント20の文書データベース選択手段202は、送られてきた順序にしたがって文書データベースを表示する。次に、利用者は文書データベース選択手段202により検索対象となる文書データベースを選択し、検索要求入力手段201に検索要求を入力する。検索要求と検索対象となる文書データベースの情報は連想検索サーバ30に送られる(T3)。

## 【0035】

連想検索サーバ30の検索要求解析手段301は検索要求を解析し、その結果を検索要求発行手段302が検索サーバ40に送る(T4)。検索サーバ40の検索手段402は文書データベース403から検索要求に関連する文書を検索し、その結果を連想検索サーバ30に送る(T5)。連想検索サーバ30の概略単語要求手段303は、送られてきた文書の概略を検索サーバ40に要求する(T6)。検索サーバ40の概略作成手段401は送られてきた文書の概略を作成し、連想検索サーバ30に送る(T7)。

## 【0036】

連想検索サーバ30の連想検索記録テーブル保持手段304は、この検索の履歴(キーワードから検索サーバ40への検索)を記録する。表示順序変更手段305は連想検索記録テーブルを参照して文書データベースの表示順序を変更し、その順序を検索クライアント20に送る。検索結果と概略単語も同時に送られる(T8)。検索クライアント20は、送られてきた文書データベースの表示順序、検索結果、概略単語を、それぞれ、文書データベース選択手段202、検索結果表示手段203、概要表示手段204で表示する。

## 【0037】

図8は、検索サーバ40を検索した結果得られた文書をキーとして検索サーバ60を検索するときの実行の流れを示す図である。まず、利用者は文書データベース選択手段202により検索対象となる文書データベースを選択し、検索結果表示手段(文書群指定手段)203を用いて記事を選択する。選択された記事と検索対象となる文書データベースの情報は連想検索サーバ30に送られる(T11)。連想検索サーバ30の概略単語要求手段303は、送られてきた文書の概略を検索サーバ40に要求する(T12)。検索サーバ40の概略作成手段401は送られてきた文書の概略を作成し、連想検索サーバ30に送る(T13)。

## 【0038】

連想検索サーバ30の検索要求発行手段302は、検索サーバ40から返された概略単語を検索要求として検索サーバ60に送る(T14)。検索サーバ60の検索手段602は文書データベース603から検索要求に関連する文書を検索



し、その結果を連想検索サーバ30に送る(T15)。連想検索サーバ30の概略単語要求手段303は、送られてきた文書の概略を検索サーバ60に要求する(T16)。検索サーバ60の概略作成手段601は送られてきた文書の概略を作成し、連想検索サーバ30に送る(T17)。

## 【0039】

連想検索サーバ30の連想検索記録テーブル保持手段304は、この検索の履歴(検索サーバ40から検索サーバ60の検索)を記録する。表示順序変更手段305は連想検索記録テーブルを参照して文書データベースの表示順序を変更し、その順序を検索クライアント20に送る。検索結果と概略単語も同時に送られる(T18)。検索クライアント20は、送られてきた文書データベースの表示順序、検索結果、概略単語を、それぞれ、文書データベース選択手段202、検索結果表示手段203、概要表示手段204で表示する。

## 【0040】

図9は、検索サーバ40を検索した結果得られた文書の概略単語をキーとして検索サーバ60を検索するときの実行の流れを示す図である。まず、利用者は文書データベース選択手段202により検索対象となる文書データベースを選択し、概要表示手段(単語群指定手段)204を用いて単語を選択する。選択された単語と検索対象となる文書データベースの情報は連想検索サーバ30に送られる(T21)。連想検索サーバ30の検索要求発行手段302は、送られてきた単語を検索要求として検索サーバ60に送る(T22)。検索サーバ60の検索手段602は文書データベース603から検索要求に関連する文書を検索し、その結果を連想検索サーバ30に送る(T23)。連想検索サーバ30の概略単語要求手段303は送られてきた文書の概略を検索サーバ60に要求する(T24)。検索サーバ60の概略作成手段601は送られてきた文書の概略を作成し、連想検索サーバ30に送る(T25)。

## 【0041】

連想検索サーバ30の連想検索記録テーブル保持手段304は、この検索の履歴(検索サーバ40から検索サーバ60の検索)を記録する。表示順序変更手段305は連想検索記録テーブルを参照して文書データベースの表示順序を変更し

、その順序を検索クライアント 2 0 に送る。検索結果と概略単語も同時に送られる (T 2 6)。検索クライアント 2 0 は、送られてきた文書データベースの表示順序、検索結果、概略単語を、それぞれ、文書データベース選択手段 2 0 2、検索結果表示手段 2 0 3、概要表示手段 2 0 4 で表示する。

#### 【 0 0 4 2 】

連想検索記録テーブルとして、連想検索サーバのすべての利用者の履歴をまとめて記録するだけでなく、利用者ごとに個別に記録することで、利用者の個々の利用履歴に応じた文書データベースの表示順序を提供することができる。利用者ごとに連想検索記録テーブルを保持する場合のシステム構成例を図 1 0 に示す。

#### 【 0 0 4 3 】

図 1 0 のシステム構成例では、図 1 の連想検索サーバ 3 0 に利用者登録手段 3 0 7 1、利用者認証手段 3 0 7 2、利用者管理手段 3 0 7 3 が追加されている。利用者登録手段 3 0 7 1 は、新規利用者を登録するための手段である。利用者認証手段 3 0 7 2 は既登録利用者の認証を行なう手段であり、例えば、利用者ごとのパスワードで認証を行なう。利用者管理手段 3 0 7 3 は、どの利用者がどの連想検索記録テーブルに対応しているかを管理する手段である。また、図 1 の検索クライアント 2 0 に利用者認証画面表示手段 2 0 5 が追加されている。利用者認証画面表示手段 2 0 5 は、利用者が連想検索サーバ 3 0 に接続する際に利用者認証画面を表示する手段であり、例えば、図 1 1 のような画面を表示する。

#### 【 0 0 4 4 】

利用者登録をした後、利用者認証をする場合の処理の流れを図 1 2 を用いて説明する。利用者は利用者名入力ボックス 2 0 5 1 に利用者名、パスワード入力ボックス 2 0 5 2 にパスワードを入力し、「登録」ボタン 2 0 5 4 をクリックする。利用者名とパスワードが連想検索サーバ 3 0 に送られ (T 3 1)、利用者登録手段 3 0 7 1 により利用者登録が行なわれる。さらに利用者管理手段 3 0 7 3 により連想検索記録テーブルとの対応がとられる。

#### 【 0 0 4 5 】

詳細を述べると、利用者ごとに連想検索記録テーブルを保持する場合、連想検索記録テーブル保持手段は、図 1 3 のように利用者別の連想検索記録テーブルを

保持する。利用者管理手段 3 0 7 3 は、図 1 4 のような利用者管理テーブルを用意する。利用者管理テーブルには、利用者ごとに、利用者 ID、利用者名、パスワード、対応する連想検索記録テーブルの ID、登録日などが記録される。利用者管理テーブルには任意の利用者情報を記録することができるが、最低限、利用者名、パスワード、テーブル ID の項目は必須である。利用者登録の場合、利用者登録手段 3 0 7 1 は、利用者管理テーブルに必要な情報を追加する。

#### 【 0 0 4 6 】

登録に成功すると検索クライアント 2 0 に処理が戻り (T 3 2)、利用者認証画面表示手段 2 0 5 により、図 1 1 の利用者認証画面が表示される。利用者は先に登録した利用者名とそれに対応するパスワードを利用者名入力ボックス 2 0 5 1、パスワード入力ボックス 2 0 5 2 に入力し、「ログイン」ボタン 2 0 5 3 をクリックすることで連想検索サーバ 3 0 に接続される (T 3 3)。利用者認証手段 3 0 7 2 は利用者管理テーブルを参照することにより利用者認証を行なう。利用者管理手段 3 0 7 3 は利用者管理テーブルを参照することによって用意されている連想検索記録テーブルとの対応をとる。

以上のように、利用者ごとに連想検索記録テーブルを保持することで、利用者の使用状況に応じて文書データベース表示順序を変更することができる。

#### 【 0 0 4 7 】

##### 【発明の効果】

本発明によると、利用者は、検索結果に応じて適切に並べ替えられた文書データベースリストから次の検索対象となるデータベースを選択できるため、より利便性の高い検索が可能となる。また、文書データベースの登録料や手数料の計算をより適切に行なうことが可能となる。

##### 【図面の簡単な説明】

##### 【図 1】

本発明の基本構成例を示す図。

##### 【図 2】

検索クライアントの例を示す図

##### 【図 3】

連想検索記録テーブルの例を示す図。

【図 4】

検索開始時点の検索クライアントを示す図。

【図 5】

連想検索記録テーブルの例を示す図。

【図 6】

文書データベースの表示順序が変更されたことを示す図。

【図 7】

キーワード検索するときの実行の流れを示す図。

【図 8】

文書をキーとして検索するときの実行の流れを示す図。

【図 9】

概略単語をキーとして検索するときの実行の流れを示す図。

【図 10】

利用者ごとの連想検索記録テーブルを保持する場合の構成を示す図。

【図 11】

利用者認証の例を示す図。

【図 12】

利用者認証の実行の流れを示す図。

【図 13】

利用者ごとの連想検索記録テーブルの例を示す図。

【図 14】

利用者管理テーブルの例を示す図。

【符号の説明】

10：通信ネットワーク

20：検索クライアント

201：検索要求入力手段

2011：キーワード入力ボックス

2012：キーワード検索指示ボタン

- 202 : 文書データベース選択手段
- 2021 : 文書データベース選択ボタン
- 2022 : 文書データベース名表示ボックス
- 203 : 検索結果表示手段 (文書群指定手段)
- 2031 : 閲覧ボタン
- 2032 : 文書選択ボタン
- 2033 : 「文書から検索」 ボタン
- 204 : 概要表示手段 (単語群指定手段)
- 2041 : 単語選択ボタン
- 2042 : 「概略単語から検索」 ボタン
- 205 : 利用者認証画面表示手段
- 2051 : 利用者名入力ボックス
- 2052 : パスワード入力ボックス
- 2053 : 「ログイン」 ボタン
- 2054 : 「登録」 ボタン
- 30 : 連想検索サーバ
- 301 : 検索要求解析手段
- 302 : 検索要求発行手段
- 303 : 概略単語要求手段
- 304 : 連想検索記録テーブル保持手段
- 305 : 表示順序変更手段
- 306 : 登録料計算手段
- 3071 : 利用者登録手段
- 3072 : 利用者認証手段
- 3073 : 利用者管理手段
- 40 : 検索サーバ
- 401 : 概略作成手段
- 402 : 検索手段
- 403 : 文書データベース

50 : 検索サーバ

501 : 概略作成手段

502 : 検索手段

503 : 文書データベース

60 : 検索サーバ

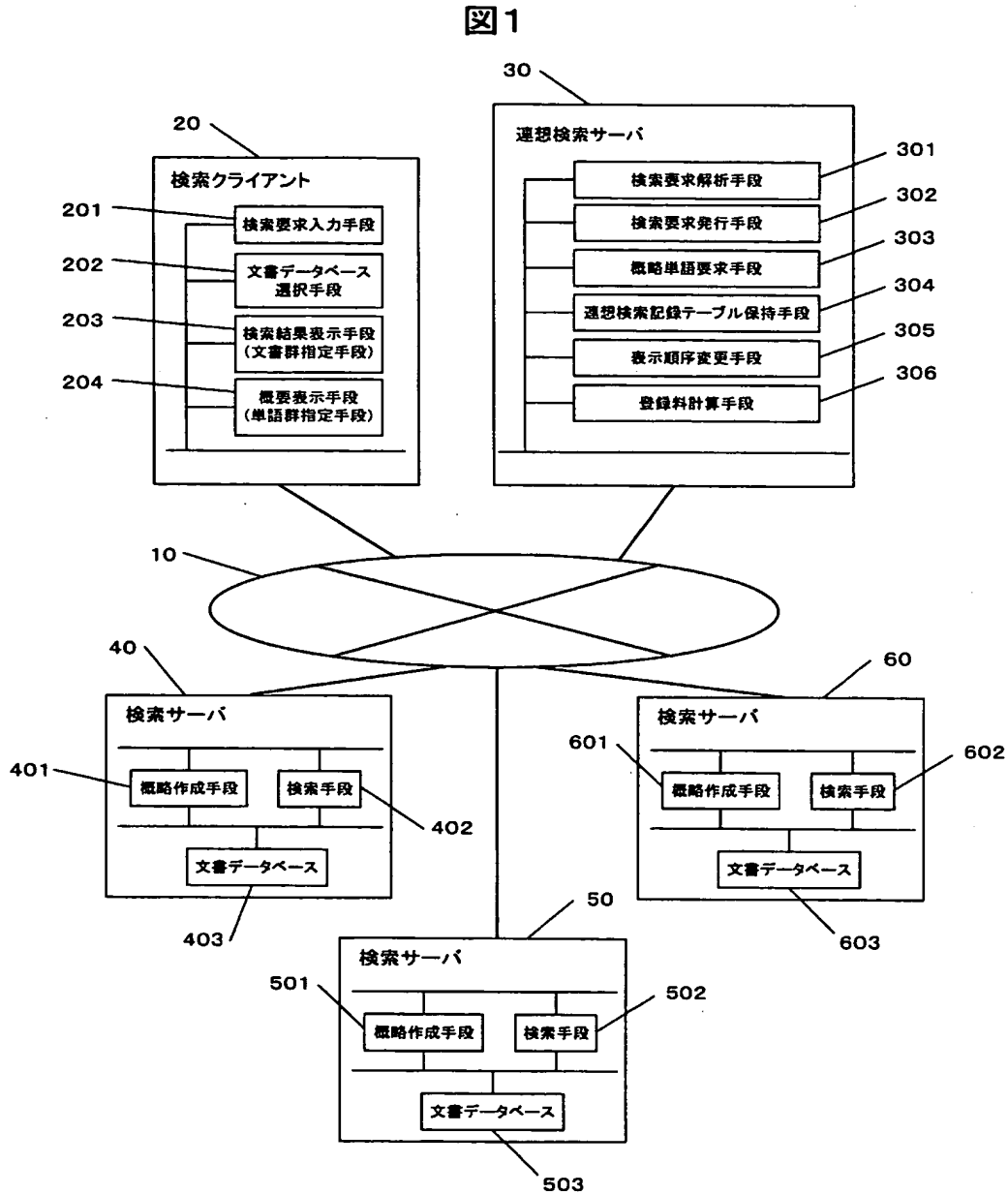
601 : 概略作成手段

602 : 検索手段

603 : 文書データベース

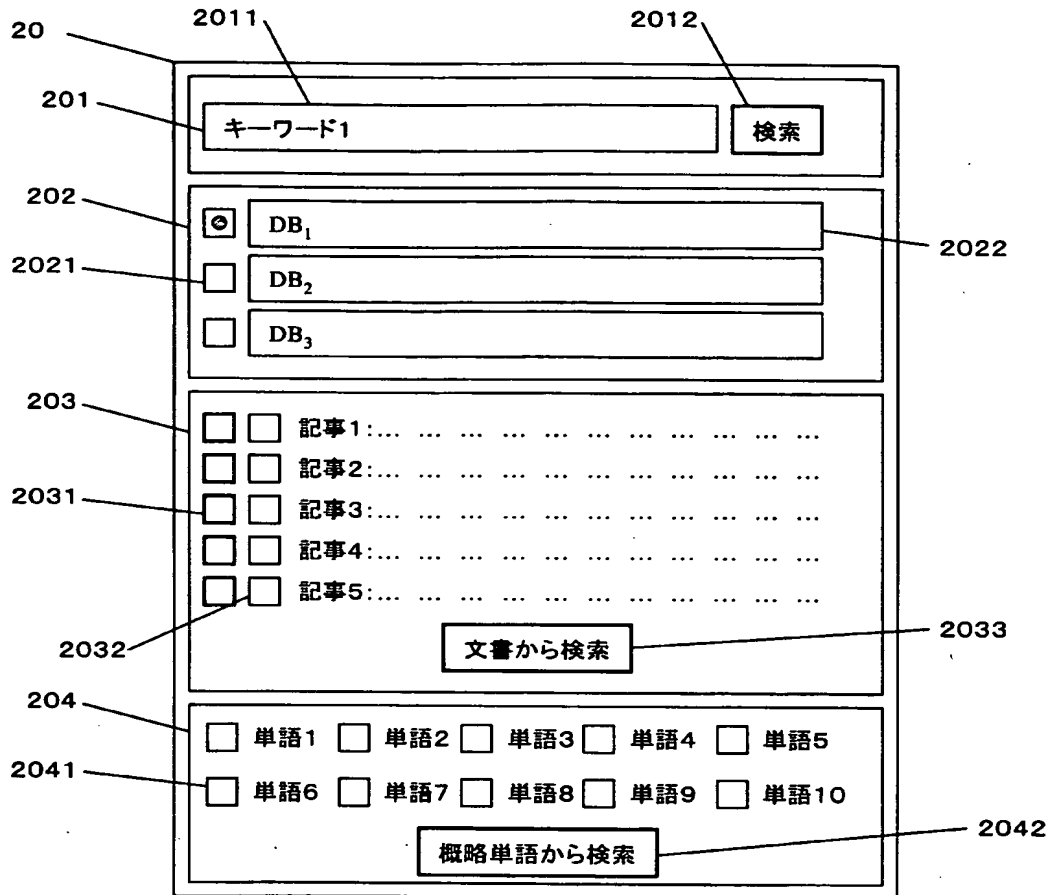
【書類名】 図面

【図 1】



【図 2】

図2





【図 3】

図 3

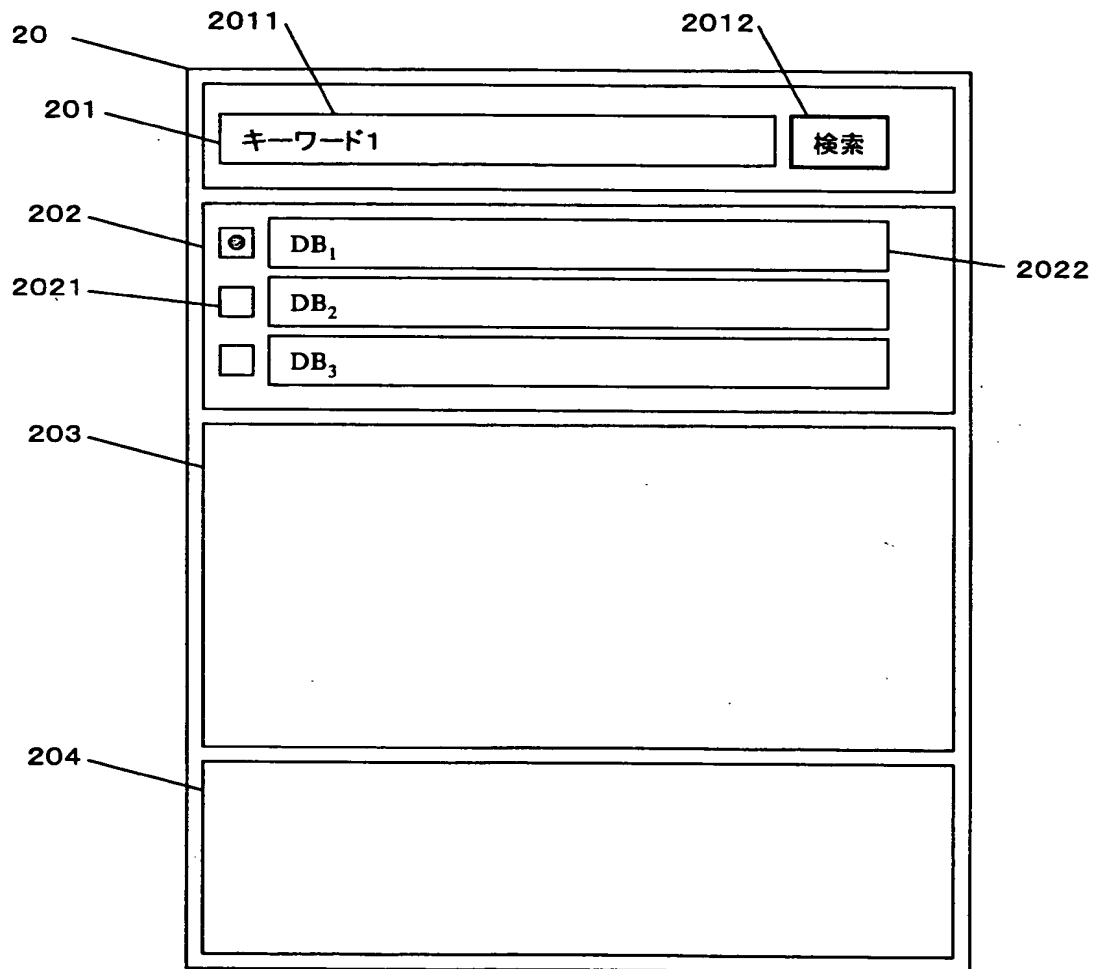
From \ To	DB <sub>1</sub>	DB <sub>2</sub>	DB <sub>3</sub>	Total
DB <sub>1</sub>	$x_{11}$	$x_{12}$	$x_{13}$	$X_1$
DB <sub>2</sub>	$x_{21}$	$x_{22}$	$x_{23}$	$X_2$
DB <sub>3</sub>	$x_{31}$	$x_{32}$	$x_{33}$	$X_3$
Keyword	$x_{41}$	$x_{42}$	$x_{43}$	$X_4$
Total	$Y_1$	$Y_2$	$Y_3$	

$$X_i = \sum_{j=1}^3 x_{ij}$$

$$Y_j = \sum_{i=1}^4 x_{ij}$$

【図 4】

図 4



【図 5】

図 5

From \ To	DB <sub>1</sub>	DB <sub>2</sub>	DB <sub>3</sub>	Total
DB <sub>1</sub>	10	20	30	60
DB <sub>2</sub>	...	...	...	...
DB <sub>3</sub>	...	...	...	...
Keyword	60	40	20	120
Total	...	...	...	

【図6】

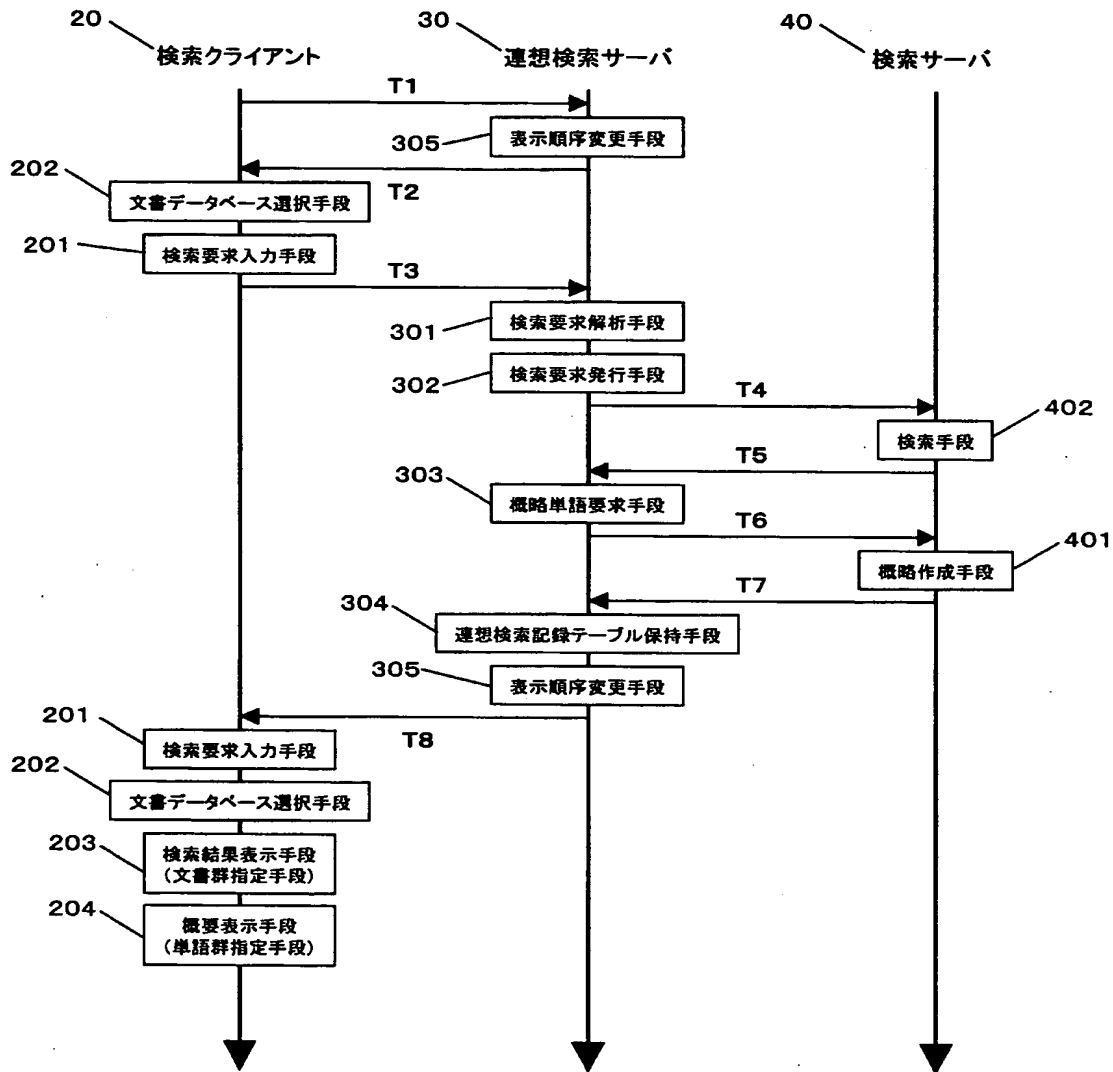
図6

The diagram illustrates a search interface with the following components and labels:

- 20**: Points to the overall search interface frame.
- 201**: Points to the keyword input field containing "キーワード1".
- 2011**: Points to the keyword input field.
- 2012**: Points to the "検索" (Search) button.
- 202**: Points to the database selection section.
- 2021**: Points to the checkboxes for database selection.
- 2022**: Points to the database labels: DB<sub>3</sub>, DB<sub>2</sub>, and DB<sub>1</sub> (which is selected).
- 203**: Points to the article selection section.
- 2031**: Points to the checkboxes for article selection.
- 2032**: Points to the article labels: 記事1, 記事2, 記事3, 記事4, and 記事5.
- 2033**: Points to the "文書から検索" (Search from documents) button.
- 204**: Points to the word selection section.
- 2041**: Points to the checkboxes for word selection.
- 2042**: Points to the "概略単語から検索" (Search from summary words) button.
- 2042**: Also points to the word labels: 単語1 through 単語10.

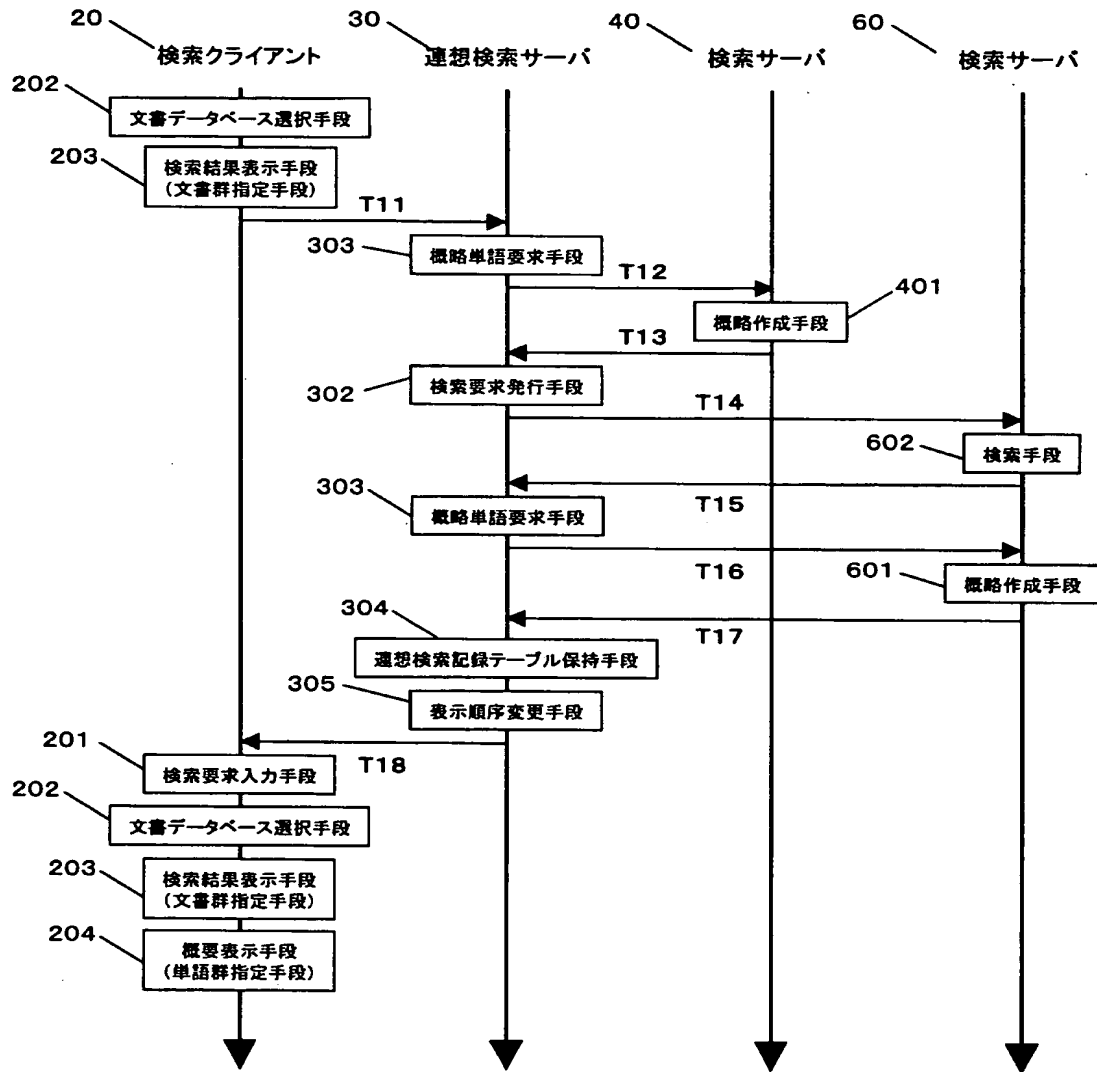
【図7】

図7



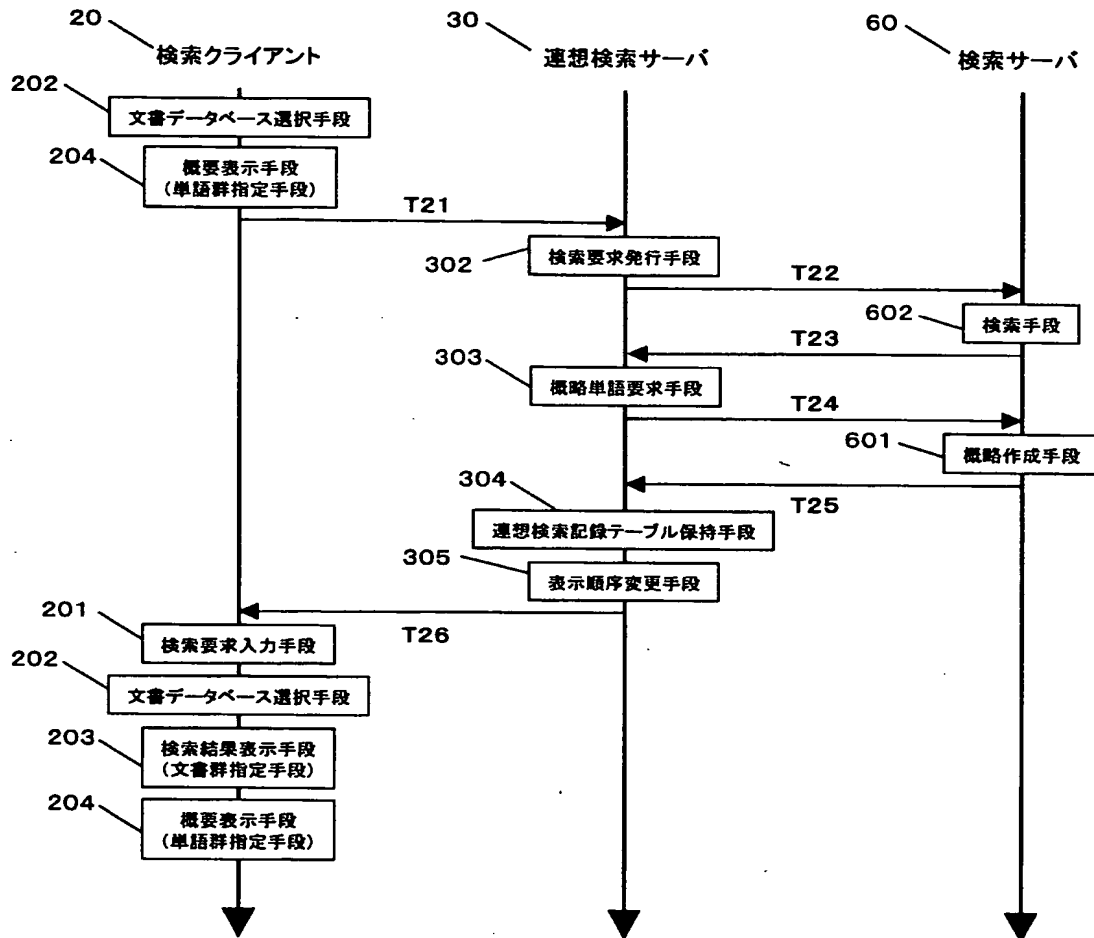
【図 8】

図 8



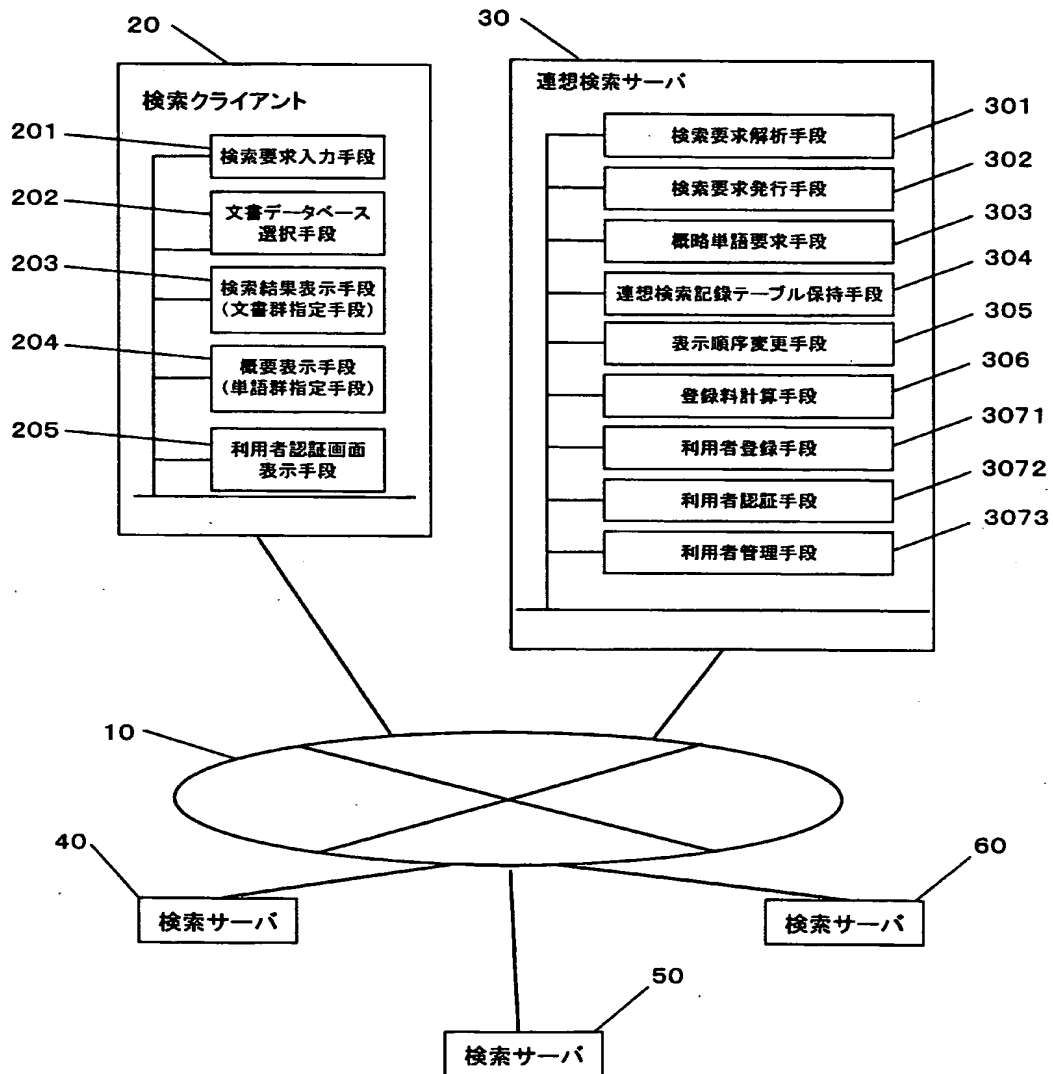
【図 9】

図9



【図10】

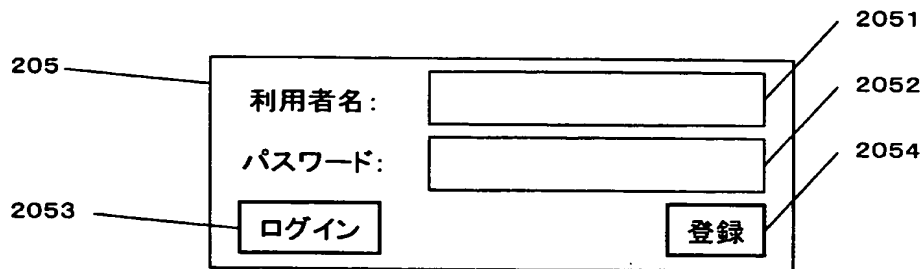
図10





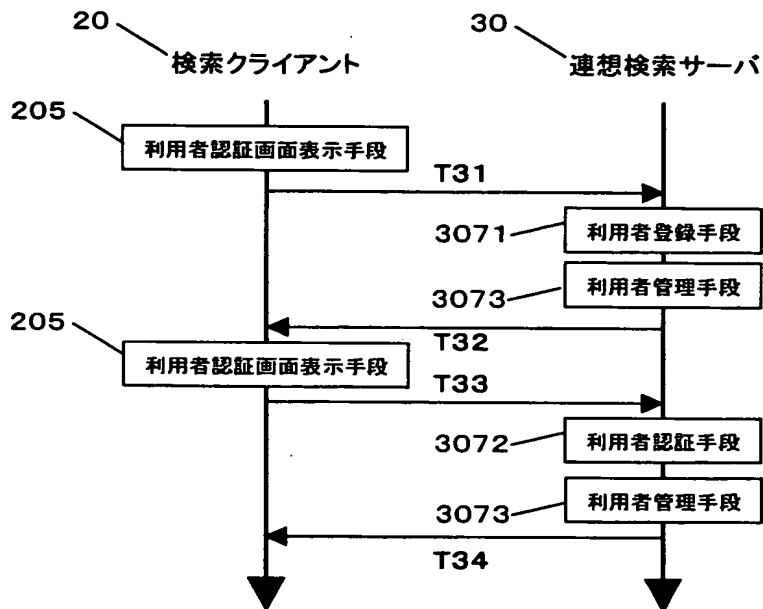
【図 11】

図 11



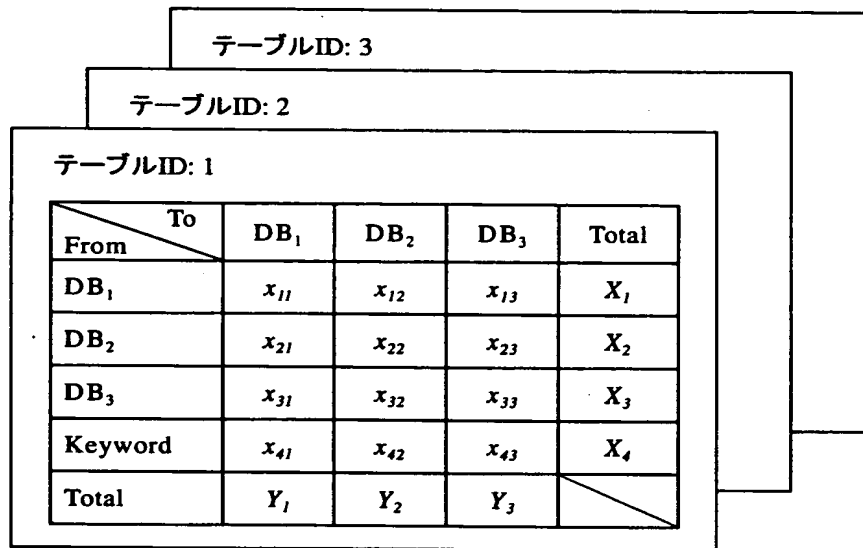
【図 12】

図 12



【図 13】

図13



【図 14】

図14

利用者ID	利用者名	パスワード	テーブルID	登録日
11	AAAA	****	1	YYYYMMDD
22	BBBB	****	2	YYYYMMDD
33	CCCC	****	3	YYYYMMDD

【書類名】 要約書

【要約】

【課題】 連想検索サーバに登録されている文書データベースの順序付けを適切に行なう。

【解決手段】 複数の文書データベース 4 0 3, 5 0 3, 6 0 3 を関連付けて連想検索できる連想検索サーバ 3 0 において、連想検索の履歴を連想検索記録テーブルとして連想検索記録テーブル保持手段 3 0 4 により保持し、それを用いて、表示手段変更手段 3 0 5 により、文書データベース選択手段 2 0 2 が提示する文書データベースの表示順序を適切にすること、あるいは、登録料計算手段 3 0 6 により、連想検索サーバに登録される文書データベース 4 0 3, 5 0 3, 6 0 3 の登録料の計算を適切に行なうことを可能とする。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005108]

1. 変更年月日	1990年 8月31日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台4丁目6番地
氏 名	株式会社日立製作所